Some notes on ARCH models

Marcin Pawlik

January 11, 2005

1 Introduction

The autoregressive conditional heteroscedasticity (ARCH) class of models has been introduced first by R.F Engle [4]. Parallel to the success of standard linear time series models, arising from the use of the conditional versus unconditional mean, the key insight offered by ARCH model lies in the distinction between conditional and unconditional second order moments. While the unconditional covariance matrix for the variables may be time invariant, the conditional variances and covariances often depend non-trivially on the past states of the world. Understanding this intertemporal dependence is crucially important for many issues in macroeconomics and finance. ARCH models are models that can represent locally non-stationary stochastic processes which are asymptotically stationary. It means that parameters determining the density function of the distribution $f_t(x)$ fluctuate in period t. Such local fluctuations do not impair the hypothesis of the existence of the asymptotic form for density function $P(x)$.

ARCH processes constitute a stable ground for stochastic models with discrete time, for which the variance in period t depends conditionally on some of previous squared disturbance terms. ARCH processes define a whole class of stochastic models while single models are characterized with various

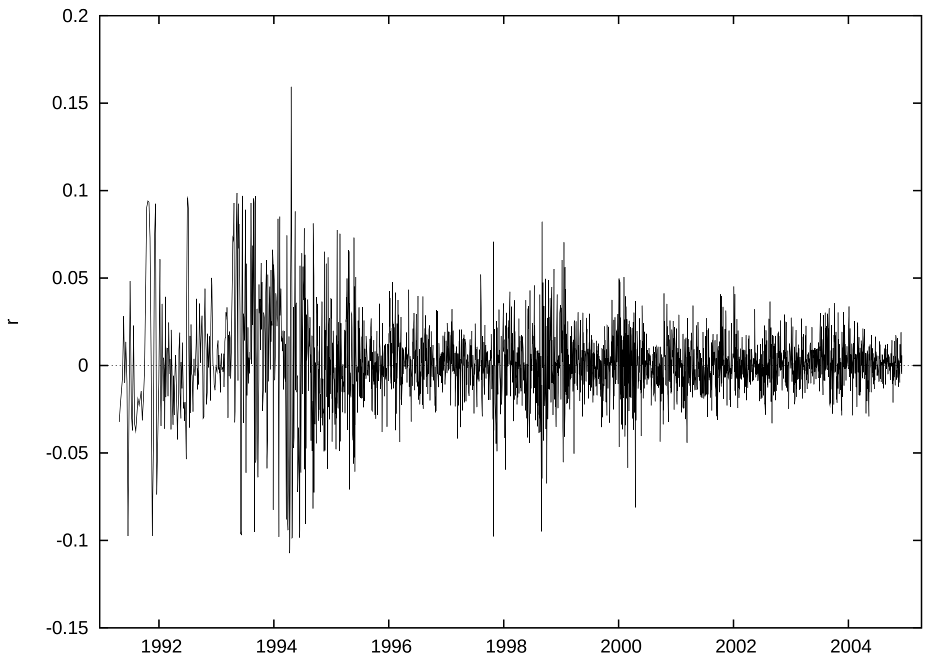

Figure 1: WIG20 rates of return 1991-2004

steering parameters and various functional forms for density functions that generate a variable in period t.

2 Volatility clustering

In many econometric research concerning financial time series the subject of modeling usually is the rate of return, usually defined in a following manner:

$$r_t = (y_t - y_{t-1}) / y_{t-1} \tag{1}$$

where y_t is the quotation of a given financial instrument. Alternatively rate of return r_t may be defined following:

$$r_t = lny_t - lny_{t-1} = ln (y_t / y_{t-1}) \tag{2}$$

The rates of return, especially those on the stock exchange have a few typical properties. One of them is volatility clustering. This means that large changes tend to be followed by large changes of either sign and small changes tend to be followed by small changes. This phenomenon is immediately apparent when assets returns are plotted through time. To illustrate that Figure 1 presents the daily capital gains on the WIG index from 1991 to 2004. Greater volatility of dependent variable results in an increase of the disturbance term variance. There have been many attempts to determine the cause of this phenomenon. The most popular theories suggest that this is a result of irregular information flow. Moreover rates of return are leptocurtic distributed which means that the probability of an unusual change is greater than it would be in case of normal distribution. Another one is so called leverage effect which refers to the tendency in stock prices to be negatively correlated with changes in stock volatility[3]. The next property are non-trading periods which leads to the situation when information accumulates while the markets are closed and is then reflected in the prices after the markets reopen. It has been showed that informations accumulates more slowly when the markets are closed. Forseable events influent high ex-ante volatility. This especially noticeable before earnings announcements or central bank statements.

3 Model ARCH(S)

The ARCH regression model is obtained by assuming that the mean of y_t is given as $x_t\beta$, a linear combination of lagged endogenous and exogenous variables included in the information set Ψ_{t-1} with β a vector of unknown parameters. Formally,

$$y_t|\Psi_t \sim N\left(x_t\beta t, h_t\right), \tag{3}$$

$$h_t = h\left(\epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_{t-p}, \alpha\right), \qquad \epsilon_t = y_t - x_t\beta. \tag{4}$$

The ARCH regression models have a variety of characterictics which make them attractive for econometric applications. Econometric forecasters have found their ability to predict the future varies from one period to another. In practice, the inherent uncertainty or randomness associated with different forecast periods seems to cary widely over time. Moreover large and small errors tend to cluster together. This analysis immediately suggests the usefulness of the ARCH model where the underlying forecast variance may change over time and is predicted by past forecast errors.

A second example can be found in monetary theory and the theory of finance. By the simplest assumptions, portfolios of financial assets are held as functions of the expected meads and variances of the rates of return. Any shifts in asset demand must be associated with changes in expected means and variances of the rates of return. If the mean is assumed to follow a standard regression or time-series model, the variance is immediately constrained to be constant over time. The use of an exogenous variable to explain changes in variance is usually not appropriate.

A third interpretation is that the ARCH regression model is an approximation to a more complex regression which has non-ARCH disturbances. The ARCH specification might then be pisking up the effect of variables omitted from the estimated model. The existence of an ARCH effect would be interpreted as evidence of misspecification, either by omitted variables or through structural change. In this is the case, ARCH may be a better approximation to reality than making standard assumptions about the disturbances, but trying to find the omitted variable or determine the nature of the structural change would be even better.

In respect of the fact that in many econometric models the heteroscedasticity cannot be eliminated even by re-specification of proper equations R.

Engle proposed to express the disturbance term variance with an AR(S) process:

$$\xi_t^2 = \gamma_0 + \sum_{s=1}^{S} \gamma_s \xi_{t-s}^2 + \eta_t, \quad \gamma_0 > 0, \gamma_s > 0 \tag{5}$$

where $\eta_t : IID\left(0, \sigma_\eta^2\right)$ and then to use this additional information in likelihood function for MLE. The key of this method is a correction of estimator and its properties. Stochastic process $\{\xi_t\}$ described by (5) is called autoregressive conditional heteroskedastic (ARCH) process. It is often denoted in the following way:

$$\xi_t = \upsilon_t \sqrt{h_t} \tag{6}$$

where $\upsilon_t : IID\,(0{,}1)\,, \xi_t : IID\left(0{,}\sigma_\xi^2\right)$ and

$$h_t = \gamma_0 + \sum_{s=1}^{S} \gamma_s \xi_{t-s}^2. \tag{7}$$

That is why it can be represented in such a manner:

$$\upsilon_t^2 h_t = h_t + \eta_t \tag{8}$$

The conditional mean and conditional variance are equal:

$$E\left(\xi_t | \xi_{t-1}\right) = 0 \tag{9}$$

$$D^2\left(\xi_t | \xi_{t-1}\right) = h_t \tag{10}$$

which means that the latter is time variant and:

$$E\left(\xi_t\right) = 0 \qquad D^2\left(\xi_t\right) = \sigma_\xi^2 = \frac{\gamma_0}{\left(1 - \sum_{s=1}^{S} \gamma_s\right)}. \tag{11}$$

Linear ARCH(S) model assuming that disturbance term can be represented by (6) and (7) can be represented as follows:

$$r_t = x_{(k)t} \alpha_{(k)} + \xi_t \tag{12}$$

$$\xi_t = \upsilon_t \sqrt{h_t} \tag{13}$$

$$h_t = \gamma_0 + \sum_{s=1}^{S} \gamma_s \xi_{t-s}^2 \tag{14}$$

for which:

$$E(r_t) = E(r_t | r_{t-1}) = x_{(k)t} \alpha_{(k)} \tag{15}$$

$$D^2 (r_t | r_{t-1}) = h_t \tag{16}$$

4 The likelihood function

The parameters of ARCH models may be estimated in OLS procedure because $x_{(k)t}$ and ξ_t are nor correlated and $\upsilon_t : IID(0{,}0)$. However the ML estimator is asymptotically more effective than the OLS one.

Let l be the average log likelihood and l_t be the log likelihood of the tth observation and T the sample size. Then

$$l = \frac{1}{T} \sum_{t=1}^{T} l_t, \qquad l_t = -\frac{1}{2} ln(2\Pi) - \frac{1}{2} ln h_t - \frac{1}{2} \xi_t^2 / h_t. \tag{17}$$

To estimate the unknown parameters α, this likelihood function can be maximized. The first order conditions are:

$$\frac{\partial l_t}{\partial \alpha} = \frac{1}{2h_t} \frac{\partial h_t}{\partial \alpha} \left(\frac{y_t^2}{h_t} - 1 \right) \tag{18}$$

and the Hessian is

$$\frac{\partial^2 l_t}{\partial \alpha \partial \alpha'} = -\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \alpha} \frac{\partial h_t}{\partial \alpha'} \left(\frac{y_t^2}{h_t} \right) + \left(\frac{y_t^2}{h_y} - 1 \right) \frac{\partial}{\partial \alpha'} \left(\frac{1}{2h_t} \frac{\partial h_t}{\partial \alpha} \right). \tag{19}$$

5 ARCH effect testing

In the linear regression model, with or without lagged-dependent variables, OLS is the appropriate procedure if the disturbances are not conditionally heteroscedastic. Because the ARCH model requires iterative procedures, it may be desirable to test whether it is appropriate before going to the effort to estimate it. The Lagrange multiplier test procedure is ideal for this as in many similar cases.

Under the null hypothesis, $\alpha_1 = \alpha_2 = \ldots = \alpha_p = 0$. The test is based upon the score under the null and the information matrix under the null. Consider the ARCH model with $h_t = h\left(z_t\alpha\right)$, where h is some differentiable function which, therefore, includes both the linear and exponential cases as well as lots of others and $z_t = \left(1, e_{t-1}^2, \ldots, e_{t-p}^2\right)$ where e_t are the ordinary least squares residuals. Under the null, h_t is a constant denoted h^0. Writing $dh_t/d\alpha = h' z_t'$, where h' is the scalar derivative of h, the score and information can be written as

$$\frac{\partial l}{\partial \alpha}|_0 = \frac{h'}{2h^0} \sum_t z_t' \left(\frac{e_t^2}{h^0} - 1\right) = \frac{h'^0}{2h^0} z' f^0, \qquad \mathcal{J}_{\alpha\alpha}^0 = \frac{1}{2}\left(\frac{h'^0}{h^0}\right)^2 E z' z \tag{20}$$

and, therefore, the LM test statistic can be consistently estimated by

$$\xi * = \frac{1}{2} f'^0 z \left(z' z\right)^{-1} f^0 \tag{21}$$

where $z' = \left(z_1', \ldots, z_T'\right)$, f^0 is the column vector of $\left(\frac{e_t^2}{h^0} - 1\right)$. An asymptotically equivalent statistic would be

$$\xi = T f'^0 z \left(z' z\right)^{-1} z' f^0 / f'^0 f^0 = TR^2 \tag{22}$$

where R^2 is the squared multiple correlation between f^0 and z. The test procedure is to run the OLS regression and save the residuals. Regress the squared residuals on a constant and p lags and test TR^2 as a χ_p^2. This will be an asymptotically locally most powerful test, a characterization it shares with likelihood ratio and Wald tests.

6 Other ARCH class models

Thanks to numerous possible specifications of the conditional variance function there are many modifications of ARCH models. Those most significant are GARCH models with their variations.

6.1 GARCH(S,Q)

In 1986 Bollerslev proposed a generalization of ARCH(S) which is called generalized autoregressive conditional heteroscedasticity (GARCH). The difference lies in the function of conditional variance that is given with equation

$$h_t = \gamma_0 + \sum_{s=1}^{S} \gamma_s \xi_{t-s}^2 + \sum_{q=1}^{Q} \phi_q h_{t-q} \tag{23}$$

where $S > 0, Q \geq 0, \gamma_0 > 0, \gamma_s \geq 0, \phi_q \geq 0$.

In GARCH(S,Q) model there have been introduced an autoregressive process in h_t. While for $Q = 0$ this process can be reduced to ARCH(S), the ARCH may be treated as specific case of GARCH. Empirical research has shown that GARCH(1,1) with following h_t

$$h_t = \gamma_0 + \gamma_1 \xi_{t-1}^2 + \phi_1 h_{t-1} \tag{24}$$

where $\hat{\gamma_1} + \hat{\phi_1} \approx 1$ is good enough to descript precisely financial phenomena. Those models for which:

$$\sum_{s=1}^{S} \gamma_s + \sum_{q}^{Q} \phi_q = 1 \tag{25}$$

are belong to the integrated in variance ARCH models (IGARCH).

6.2 NARCH(S,Q)

A modification of GARCH model is a nonlinear ARCH - NARCH with following function of conditional variance:

$$h_t^{\mu/2} = \gamma_0 + \sum_{s=1}^{S} \gamma_s |\xi_{t-s}|^{\mu} + \sum_{q=1}^{Q} \phi_q h_{t-q}^{\mu/2}, \tag{26}$$

where $S > 0$, $Q \geq 0$, $\gamma_0 > 0$, $\gamma_s \geq 0$, $\phi_q \geq 0$ and $\mu \in \mathbf{N}$.

For $\mu = 2$ NARCH(S,Q) is reduced to GARCH(S,Q) and for $\mu = 1$ to:

$$\sqrt{h_t} = \gamma_0 + \sum_{s=1}^{S} \gamma_s |\xi_{t-s}| + \sum_{q=1}^{Q} \phi_q \sqrt{h_{t-q}}. \tag{27}$$

6.3 EGARCH(S,Q) and MARCH(S,Q)

GARCH(S,Q) and NARCH(S,Q) models allow to include the volatility clustering but they have some drawbacks as well. Firstly, the conditional variance is independent from the sign of ξ_t. On real markets, however, rates of return are negatively correlated with their volatility. The variance of shares prices usually increases as a result of a negative information and decreases in an opposite situation. Secondly, GARCH parameters can be only positive in order to assure positive sign of h_t. This means that an increase in the value of ξ_t always results in an increase of the conditional variance, which excludes oscillation of h_t. Lastly, the effects of a shock may not be dying out but rather cyclical.

Exponential GARCH, that is EGARCH(S), does not have that drawbacks. Its conditional variance function takes the following form:

$$log h_t = \gamma_0 + \sum_{s=1}^{S} \gamma_s \left(\delta_1 \upsilon_{t-s} + \delta_2 \left(|\upsilon_{t-s}| - \sqrt{2/\pi} \right) \right) \tag{28}$$

where $S > 0$. The same properties has EGARCH(S,Q):

$$log h_t = \gamma_0 + \sum_{s=1}^{S} \gamma_s \left(\delta_1 \upsilon_{t-s} + \delta_2 \left(|\upsilon_{t-s}| - \sqrt{2/\pi} \right) \right) + \sum_{q=1}^{Q} \phi_q log h_{t-q} \tag{29}$$

where $S > 0$, $Q > 0$ and $\upsilon_t = \xi_t / \sqrt{h_t}$.

Another one GARCH modification is multiplicative ARCH (MARCH) with a conditional variance function:

$$log h_t = \gamma_0 + \sum_{s=1}^{S} \gamma_s log \upsilon_{t-s}^2 + \sum_{q=1}^{Q} \phi_q \left(log \upsilon_{t-q}^2 - log h_{t-q} \right) \tag{30}$$

where $S > 0$, $Q \geq 0$. These models are nonlinear ARCH models.

6.4 QARCH(S,Q) and AARCH(S,Q)

The asymmetry of h_t is taken into account by QARCH modification:

$$h_t^{\mu/2} = \gamma_0 + \sum_{s=1}^{S} \gamma_s |\xi_{t-s} - \kappa|^{\mu} + \sum_{q=1}^{Q} \phi_q h_{t-q}^{\mu/2}, \tag{31}$$

where $S > 0$, $Q \geq 0$, $\gamma_0 > 0$, $\gamma_s \geq 0$, $\phi_q \geq 0$ and $\mu \in \mathbf{N}$, $\kappa \neq 0$. A specific case of QARCH is asymmetric ARCH (AARCH):

$$h_t = \gamma_0 + \gamma_1 \xi_{t-1}^2 + \delta \xi_{t-1} + \phi_1 h_{t-1}. \tag{32}$$

In this model the negative value of δ means that positive rates of return increase the value of conditional variance function more than the negative ones.

6.5 ARCH-M(S)

Within the development of ARCH methodology it has been postulated to include the variance equation into the main regression. This has been explained by the trade-off between the risk and return on investment. This proposal is brought in life by ARCH-in-Mean (ARCH-M(S)), where the conditional variance determines the value of the dependent variable.

$$r_t = x_{(k)t} \alpha_{(k)} + \psi \sqrt{h_t} + \xi_t \tag{33}$$

$$\xi_t = \upsilon_t \sqrt{h_t} \tag{34}$$

$$h_t = \gamma_0 + \sum_{s=1}^{S} \gamma_s \xi_{t-s}^2. \tag{35}$$

In application sometimes simplified ARCH-M(S) is being used:

$$r_t = \psi \sqrt{h_t} + \xi_t \quad \text{or} \quad r_t = \psi log \sqrt{h_t} + \xi_t. \tag{36}$$

ARCH-M(S) is commonly used in analysis of CAPM (Capital Assets Pricing Model) and APT (Arbitrage Pricing Theory) models.

7 Classical conformance measures

Let Ψ_{t-1} denote the information set in period $t-1$. For a good model we have:

$$E\left(r_t^2|\Psi_{t-1}\right) = E\left(\xi_t^2|\Psi_{t-1}\right) = h_t \tag{37}$$

which leads to a conclusion that the proper conformance measure for an ARCH model is the R^2 of the following regression:

$$r_t^2 = \phi_0 + \phi_1 h_t + \xi_t. \tag{38}$$

For data of higher (b) than daily frequency $r_t = lny_t - lny_{t-1/b}$ for $t = 1/b, \ldots, b/b$ and the proper statistic is R^2 of the equation

$$r_{t+1/b}^2 = \phi_0 + \phi_1 h_{t+1/b} + \xi_{t+1/b}. \tag{39}$$

Empirical research showed that R^2 for this equation is relatively low and seldom exceeds 0,05. That is as the conformance measure has been proposed[1] R^2 of following regression

$$\sum_{a=1}^{b} r_{t+a/b}^2 = \phi_0 + \phi_1 h_{t+1} + \xi_{t+1}. \tag{40}$$

This change leads to an increase in R^2. For daily data it is still about 0.05, but for hourly it increases up to 0.2-0.3 and for high frequency data (5 minutes) it reaches 0.5.

8 Direction conformance measures

The success of an investment strategy often depends not on the precise forecast of shares prices but on the ability to predict the direction in which the price flows. A measure that can reckon forecasting properties of a model is direction conformance measure

$$Q1 = \frac{N\left(\{r_t\hat{r}_t > 0\}\right)}{N\left(\{r_t\hat{r}_t \neq 0\}\right)} \tag{41}$$

where:

$\hat{r}_t$	theoretical values of the dependent variable,
$N(\{r_t\hat{r}_t > 0\})$	number of observations for which $r_t\hat{r}_t > 0$,
$N(\{r_t\hat{r}_t \neq 0\})$	number of observations for which $r_t\hat{r}_t \neq 0$.

In the context of short term investment strategies the ability to predict the turn points may be essential. This can be measured with:

$$Q2 = \frac{N\left(\{r_t\hat{r}_t > 0 | r_{t-1}r_t < 0\}\right)}{N\left(\{r_t\hat{r}_t \neq 0 | r_{t-1}r_t < 0\}\right)} \tag{42}$$

where $N\left(\{r_t\hat{r}_t > 0 | r_{t-1}r_t < 0\}\right)$ is the number of observations for which $r_t\hat{r}_t > 0$ under the condition that $r_{t-1}r_t < 0$.

On the market where there are costs like commission for every transaction corrected measures may be applied (if the commission amounts to 1% of the transaction value

$$Q1_{1\%} = \frac{N\left(\{r_t\hat{r}_t > 0 | r_t > 1\%\}\right)}{N\left(\{r_t\hat{r}_t \neq 0 | r_t > 1\%\}\right)}, \tag{43}$$

$$Q2_{1\%} = \frac{N\left(\{r_t\hat{r}_t > 0 | \left(r_{t-1}r_t < 0 \cap r_t > 1\%\right)\}\right)}{N\left(\{r_t\hat{r}_t \neq 0 | \left(r_{t-1}r_t < 0 \cap |r_t| > 1\%\right)\}\right)}. \tag{44}$$

References

[1] T. Andersen, T. Bollerslev: *Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts*, Proceedings vol. 1, Second International Conference on High Frequency Data in Finance, Olsen and Associates, Zurich 1998.

[2] T. Bollerslev, R.F. Engle, D.B. Nelson: *ARCH models*, "Handbook of Econometrics Vol. 4", Elsevier Science, 1994.

[3] J. Brzeszczyński, R. Kelm: *Ekonometryczne modele rynków finansowych*, WIG-Press, Warszawa 2002.

[4] R.F. Engle: *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*, Econometrica, Vol. 50, No. 4 (July, 1982).

[5] R.N. Mantegna, H.E Stanley: *An Introduction to Econophysics. Correlations and Complexity in Finance*, Cambridge University Press, 2000.